Cross-Correlations and their Applications to Brain Patterns and Sound Waves

Gilad Margalit and Josh Kable

Introduction

Cross-correlations are among the simplest and most broadly useful tools in signal processing, providing a quantitative measure of the similarity of two waveforms. In this paper, we show that the cross-correlation can be expressed as a vector product in Fourier series parameter space. This causes Fourier series approximations to be a precise means of comparing sets of signals, especially when those signals have relatively few data points. We apply this method to a study on oxygenation signals in the brain, where this is exactly the case. Due to this method's applicability with shorter signals, we were also able to use it to compare sound samples on the order of 10 ms in order to correlate instrument sounds, both with each other and with themselves at different points in a longer sample.

The Cross-Correlation Function

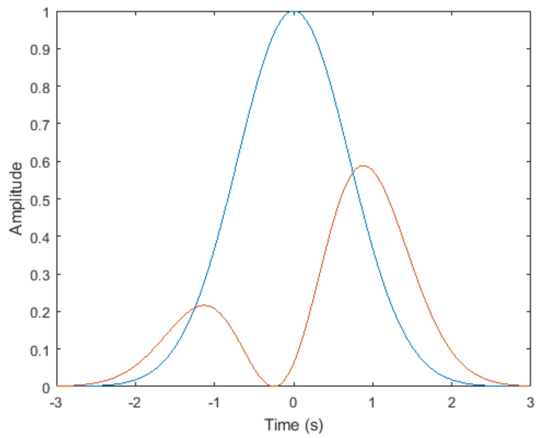
Given waveforms f(t) and g(t), the cross correlation function is defined as

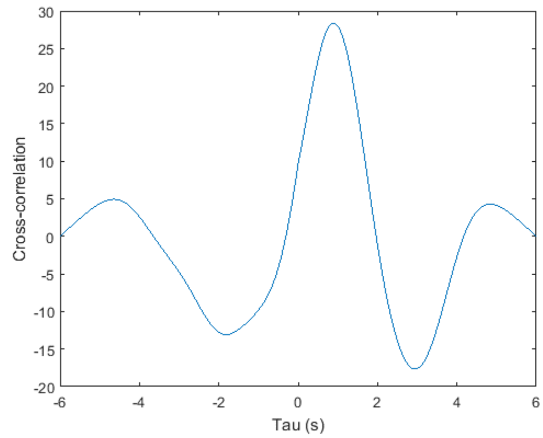
$$(f \circ g)(\tau) \equiv \int_{-\infty}^{\infty} f^*(t)g(t+\tau)dt$$

where * refers to complex conjugation and τ is the time lag between the two signals. For example, suppose we have

$$f(t) = (t+0.25)^2 e^{-t^2}$$

$$g(t) = e^{-t^2}.$$





In the plot on the left, f(t) is the orange curve, and g(t) is the blue curve. The cross-correlation function of f and g is shown on the right.

The cross-correlation function reaches its maximum value at approximately $\tau = 1$ which is defined as the amount of time that g(t) lags behind f(t). This time lag corresponds to the peak of g(t) lining up with the maximum of f(t). While f(t) has a second maximum at t = -1.15, it is a smaller peak. This example provides an intuitive understanding of how the cross-correlation function works: The cross-correlation function shifts one signal until the maxima of both signals are at the same point in time. The amount that the first signal is shifted is said to be by how much the first signal lags behind the second.

When the cross-correlation function is negative, this corresponds to a negative correlation or anti-correlation between the two signals. For our example the magnitude of the cross-correlation function at its minimum (located at $\tau = 3$) is smaller than the magnitude at its maximum. Hence we would say that the two signals are correlated rather than anti-correlated.

Supposes we have three signals a, b, and c. If we want to compare the correlation of signals a and b to correlation of signals a and c, we first need to normalize the cross-correlation function such that the cross-correlation of each signal with itself is 1. For example, suppose a and b are the same function and that c is twice that function, then, without normalization, the cross-correlation function would say that a and c are twice as correlated as a and b. Since we expect the cross-correlation of all three signals to be the same, the unnormalized cross-correlation function is somewhat dubious.

If f and g are discrete one dimensional arrays, then the integral in the cross-correlation function becomes a sum, and the cross-correlation function is a sliding dot product.

$$(f \circ g)[n] \equiv \sum_{m=-\infty}^{\infty} f^*[m] g[m+n].$$

The Fourier Series

It is a well-known property that sines and cosines obey the orthogonality relations

$$\int \sin m\omega t \sin n\omega t \, dt = \int \cos m\omega t \cos n\omega t \, dt = \pi\delta_{mn}$$

$$\int \sin m\omega t \cos n\omega t \, dt = \int \sin m\omega t \, dt = \int \cos m\omega t \, dt = 0$$

where the integrals are taken over an interval given by $\frac{2\pi}{\omega}$ which is the period when m and n are equal to one. Because of these orthogonality relations we can define a completeness relation such that an arbitrary function f(t) can be written as

$$f(t) = \frac{a_o}{2} + \sum_{n=1}^{\infty} a_n \cos n\omega t + b_n \sin n\omega t$$

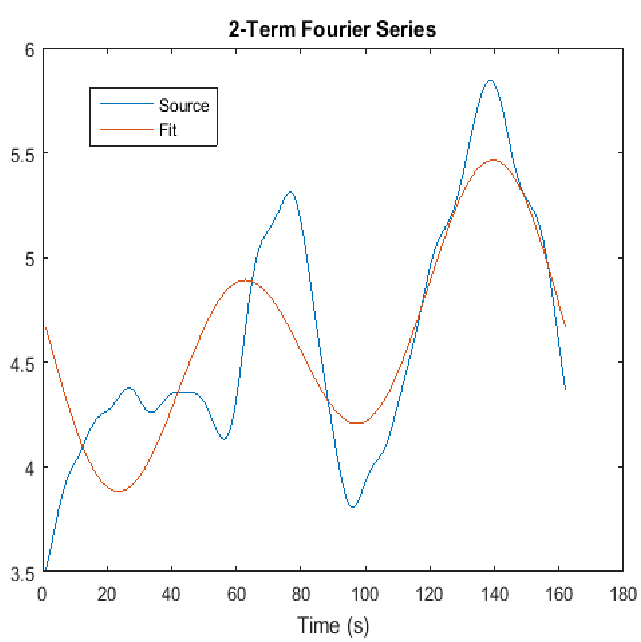
such that

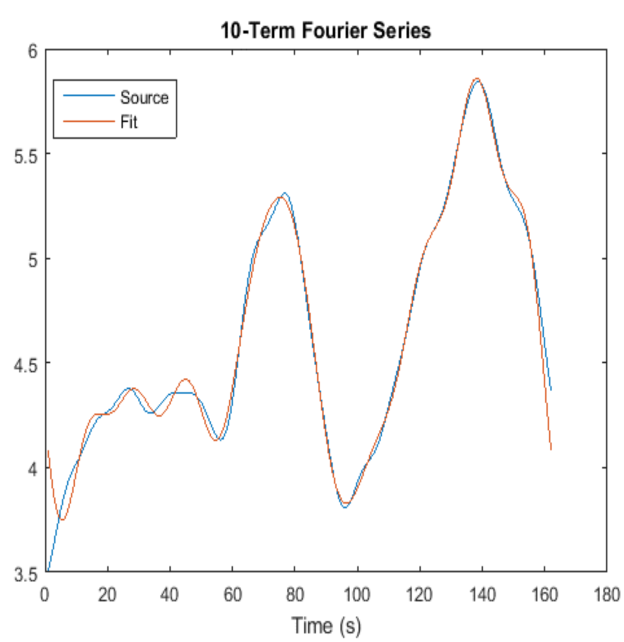
$$a_o = \frac{\omega}{\pi} \int f(t) dt$$

$$a_n = \frac{\omega}{\pi} \int f(t) \cos nwt \, dt$$

$$b_n = \frac{\omega}{\pi} \int f(t) \sin nwt \, dt.$$

Upon initial glance, it appears that we have made f(t) much more complicated; however, suppose we are given an arbitrary signal as a function of time. We can approximate this signal with a finite number of terms in the Fourier series, where it should be understood that adding more terms to the series will make the approximation more accurate.





In the figures above, the source signal is dominated by low frequency modes, so we can reasonably fit the source signal with relatively few terms in the Fourier series. In fact, this is a good way of thinking about the Fourier series. Essentially, each component of the Fourier series corresponds to a sinusoid oscillating at a particular frequency. Therefore, the relative importance of a particular frequency to the signal is encoded in the relative size of the Fourier coefficient. The a_o coefficient is the average of the source signal.

The construction of the Fourier series is a linear process which implies that we can encode the Fourier components in a linear vector space. If we approximate the source signal with k-terms in the Fourier series, then the linear vector space will have dimension 2k+1. For our purposes, we will subtract off the mean of the source signals, so $a_o = 0$, and our vector space is of dimension 2k. A vector in this space would be written $f=(a_1, a_2, \ldots a_k, b_1, b_2, \ldots b_k)$. We can promote our vector space to an inner product space by defining an inner or dot product given by

$$\langle f, g\rangle = \sum_{n=1}^{k}\left(a_{fn}a_{gn} + b_{fn}b_{gn}\right)$$

which is the conventional definition of the dot product. Since we have a definition of an inner product, we can normalize a signal such that its inner product with itself is 1:

$$\hat{f} = \frac{f}{\sqrt{\langle f, f\rangle}}$$

The hat is used to denote the fact that we now have a unit vector in this 2k dimensional space. Because the signal is now represented as a unit vector, we can understand how important each oscillation frequency is to the signal. For instance if $a_j = 0.8$, this means that $\cos(j\omega t)$ comprises $a_j^2 = 0.8^2=0.64$ or 64% of the signal.

Suppose we have two signals f(t) and g(t) which we write in Fourier series vector space as unit vectors. We can compare how similar the two signals are to each other via the inner product

$$r \equiv \langle \hat{f}, \hat{g}\rangle = \cos\theta$$

which must be bounded from negative one to one. If r = 1, then the two signals are identical up to some positive multiplicative factor. Similarly, if r = -1, then the two signals are identical up to a negative multiplicative factor. For r = 0, the two signals are orthogonal which implies that there is no correlation between them. Positive values of r indicate a correlation between the two signals, and negative values of r indicate an anti-correlation between the two signals. The larger the magnitude of r, the more strongly correlated the two signals are to each other. Hence r can be used as a measure of how correlated the two signals are to one another.

For an intuitive understanding of what the inner product means, we can decompose $\hat{f}$ into two components: one that is parallel and one that is orthogonal to $\hat{g}$. Hence,

$$\hat{f} = \langle \hat{f}, \hat{g}\rangle\hat{g} + \sqrt{1 - \langle \hat{f}, \hat{g}\rangle^2}\,\hat{g}_\perp = r\hat{g} + \sqrt{1 - r^2}\,\hat{g}_\perp$$

where $\hat{g}_\perp$ is the part of $\hat{f}$ orthogonal to $\hat{g}$. Using the second definition of r in terms of θ, we find that

$$\hat{f} = \|\hat{f}\|\cos\theta\hat{g} + \|\hat{f}\|\sin\theta\,\hat{g}_\perp = \cos\theta\hat{g} + \sin\theta\,\hat{g}_\perp\,.$$

Suppose you are given a correlation value r between $\hat{f}$ and $\hat{g}$, how similar are the two signals? The value of r encodes how much the signals overlap. Another way of thinking about the correlation value r, is that the value r^2 is the percentage of signal $\hat{f}$ which is comprised of signal $\hat{g}$. For instance, if r = 0.6, then there is an $r^2 = 0.36$ or 36% of $\hat{f}$ is comprised of $\hat{g}$.

The Fourier Series and the Cross-Correlation

In the previous section, we outlined a second method for measuring how similar two signals are to one another. In this section, we will show that the Fourier series method is actually just the cross-correlation method in disguise. Given two signals f(t) and g(t), the cross-correlation of f and g is

$$(f \circ g)(\tau) \equiv \int_{-\infty}^{\infty} f^*(t)g(t+\tau)dt.$$

We can rewrite the functions in the integrand using their Fourier series. Thus we find that

$$(f \circ g)(\tau) = \int\left[\sum_{n=1}^{k} a_{fn}\cos n\omega t + b_{fn}\sin n\omega t\right]\left[\sum_{m=1}^{k} a_{gm}\cos n\omega(t+\tau) + b_{gm}\sin n\omega(t+\tau)\right]dt$$

where we have subtracted off the average value from both signals and the integration range is of length $\frac{2\pi}{\omega}$. If we flip the order of integration and summation, the orthogonality conditions for sines and cosines reduce the expression to

$$(f \circ g)(0) = \pi\sum_{n=1}^{k}\left(a_{fn}a_{gn} + b_{fn}b_{gn}\right) = \pi\langle f, g\rangle.$$

Therefore, the cross-correlation function at τ = 0 is proportional to the Fourier series inner product. This proportionality constant is eliminated when the signals are normalized. An important note is that we have shown that this is true at τ = 0. To show that the relation holds for all τ, we need to define how the Fourier series is affected by adding τ. In particular, the Fourier coefficients at order m are

$$a_{gm}\cos n\omega(t+\tau) + b_{gm}\sin n\omega(t+\tau).$$

Now, cosine and sine follow a set of properties called the sum and difference relations which say that the previous expression is equivalent to

$a_{gm}{}'\cos n\omega(t+\tau) + b_{gm}{}'\sin n\omega(t+\tau)$ where

$a'_{gm}(\tau) = a_{gm} \cos n\omega\tau + b_{gm} \sin n\omega\tau$ and

$b'_{gm}(\tau) = b_{gm} \cos n\omega\tau - a_{gm} \sin n\omega\tau$

which means that we simply perform a Euclidean rotation by an angle $n\omega\tau$, which can be seen in the matrix equation

$$\begin{pmatrix} a_{gm}'(\tau) \\ b_{gm}'(\tau) \end{pmatrix} = \begin{pmatrix} \cos n\omega\tau & \sin n\omega\tau \\ -\sin n\omega\tau & \cos n\omega\tau \end{pmatrix} \begin{pmatrix} a_{gm} \\ b_{gm} \end{pmatrix}.$$

Hence, we may write that

$$g(t+\tau) = \sum_{m=1}^{k} a_{gm}'(\tau) \cos n\omega t + b_{gm}'(\tau) \sin n\omega t.$$

Writing the Fourier series in this way allows us to calculate the Fourier coefficients once and then simply rotate them using the aforementioned rotation matrix to determine the coefficients as a function of time delay τ. Note that only terms of the same order can rotate into each other. Therefore, $a_{gm}'(\tau)$ will only depend on terms of order m. Therefore we find that

$$(f \circ g)(\tau) = \int \left[\sum_{n=1}^{k} a_{fn} \cos n\omega t + b_{fn} \sin n\omega t\right]\left[\sum_{m=1}^{k} a_{gm}'(\tau) \cos n\omega t + b_{gm}'(\tau) \sin n\omega t\right] dt$$

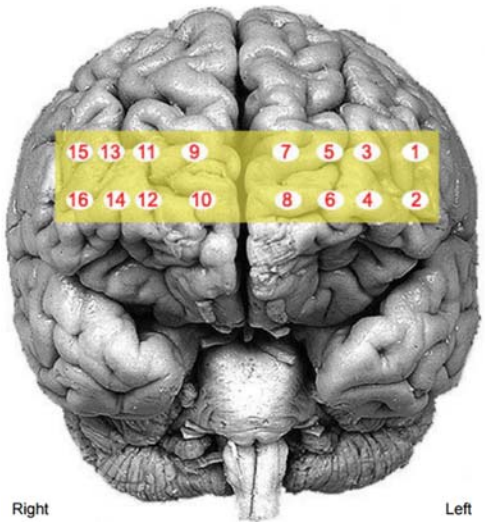
which is equal to

$$(f \circ g)(\tau) = \pi \sum_{n=1}^{k} \left(a_{fn} a_{gn}'(\tau) + b_{fn} b_{gn}'(\tau)\right) = \pi \langle f, g \rangle(\tau).$$

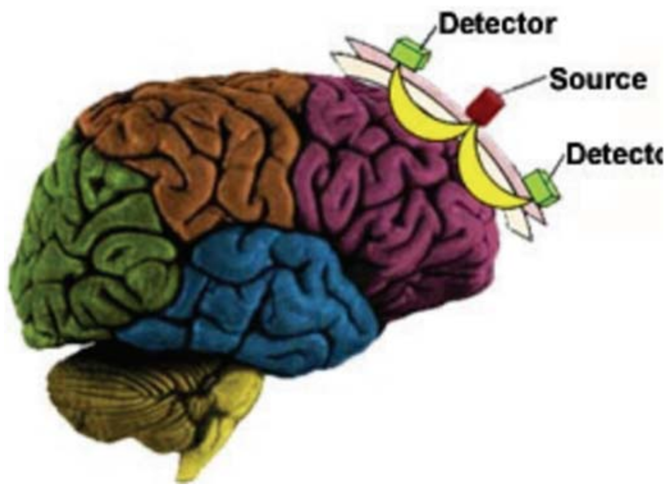
Therefore cross-correlation of f and g is proportional to the inner product of the Fourier series expansions of f and g as a function of time lag. Once we normalize our Fourier space vectors f and g we find that the two methods produce identical results.

Application to Neuroscience

A study was conducted where the oxygenation levels of hemoglobin in the frontal lobes was measured for children with Fetal Alcohol Syndrome (FAS) as well as a control group. The oxygenation in the frontal lobe is important because it oxygen is necessary to metabolize glucose which provides energy for neurons. Oxygen is transported to this area via oxygenated hemoglobin in blood. Therefore, the amount of oxygen in hemoglobin in the frontal lobes correlates how much brain activity is occurring in that region. The oxygenation levels were measured by a device with 16 optodes arranged across the frontal lobe as shown in the figure below.

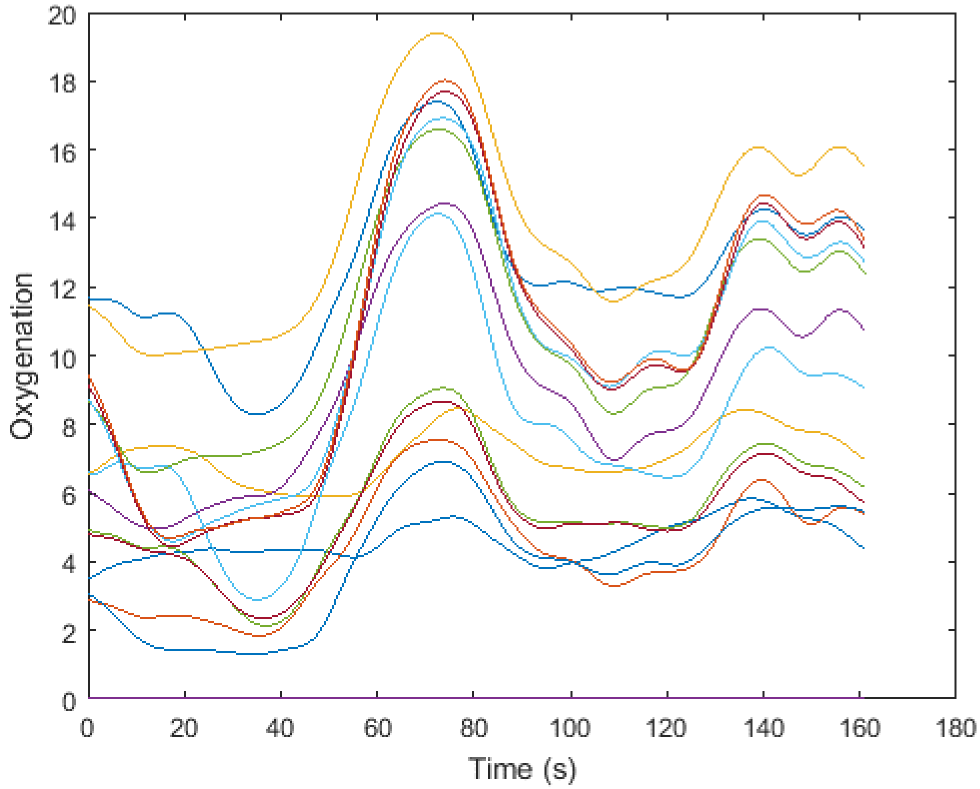


To measure the amount of oxygen in the hemoglobin, the device shines light at a given intensity on a region of the frontal lobe where the light scatters back to one of the optodes where a detector measures the new intensity of the light. Given the ratio of the intensity of the source light to the intensity of the detected light, the device calculates the amount of oxygen using the Beer – Lambert law. A cartoon description of this process is shown in the figure below.



The oxygenation levels are then recorded as a function of time with measurements taken at one second intervals. While the measurement is taken, the subjects play a computer game in which there are seven parts or blocks. During the odd number blocks, the subject wins the game, and during the even number blocks the subject loses the game. In this way, we can quantitatively examine how the subject's brain reacts to positive and negative stimuli. For this report, we examine only the win trials or positive stimuli.

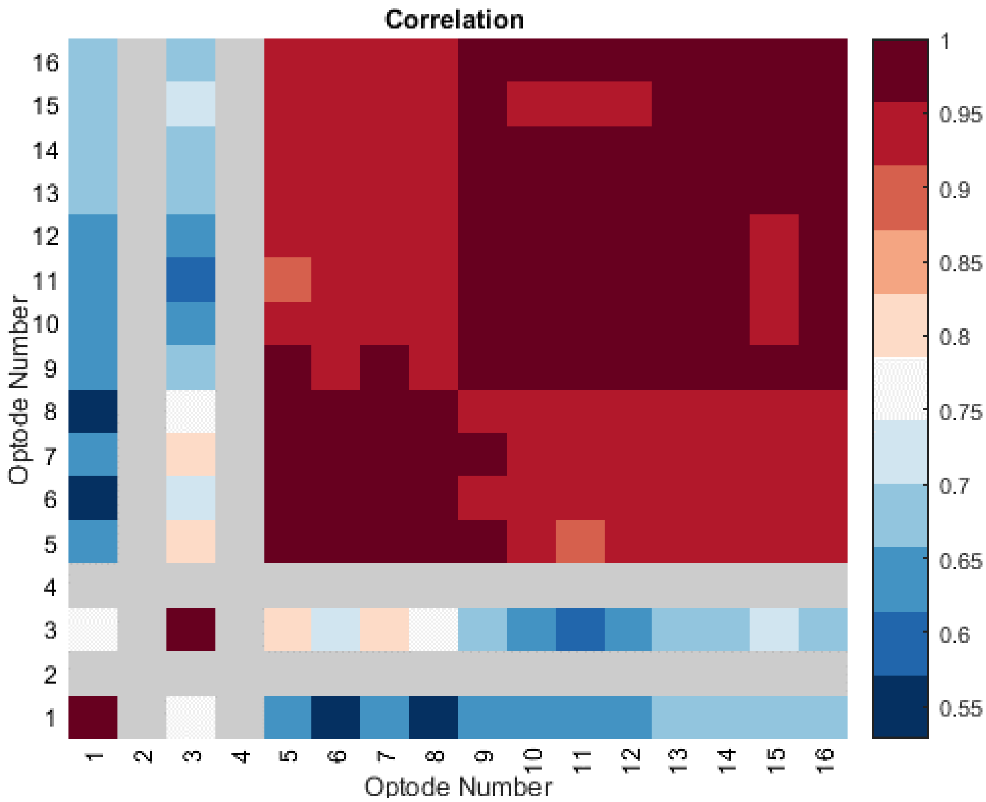
For example, the oxygenation levels of hemoglobin in the frontal lobe as a function of time of a control subject during a win trial are provided below.



From the plot, we can qualitatively see that most of the sixteen signals have a similar peak between times of sixty and eighty seconds followed by a double humped peak between 140 and 160 seconds. Overall, the signals seem qualitatively similar. We want to measure how correlated the signals are as well as how much time delay there is between two signals. To find the time delays, the simplest technique would be to apply the cross-correlation method between all combinations of the sixteen signals. This would generate 256 correlation values, which need to be normalized otherwise we cannot compare them, as well as 256 time delays. The problem with the standard cross-correlation method, in this case, is that the data was taken at one second intervals which means that we can only find time delays at one second intervals leading to errors on the order of one second caused by rounding. In turn, this limits how accurate the correlation values can be leading to even more error.

Alternatively, we can use the Fourier series inner product method which we were able to show is equivalent to the cross-correlation method. Because moving the Fourier series forward or backwards in time is achieved by performing a Euclidean rotation by an angle $n\omega\tau$, we see that for arbitrarily small τ, we can define the Fourier series. Theoretically, this implies that we can calculate arbitrarily small time delays; however, computationally, there is a limit on the accuracy based on how many digits the computer stores during each operation. For this reason, we resolve only time delays to a tenth of second which have errors on the order of a tenth of a second.

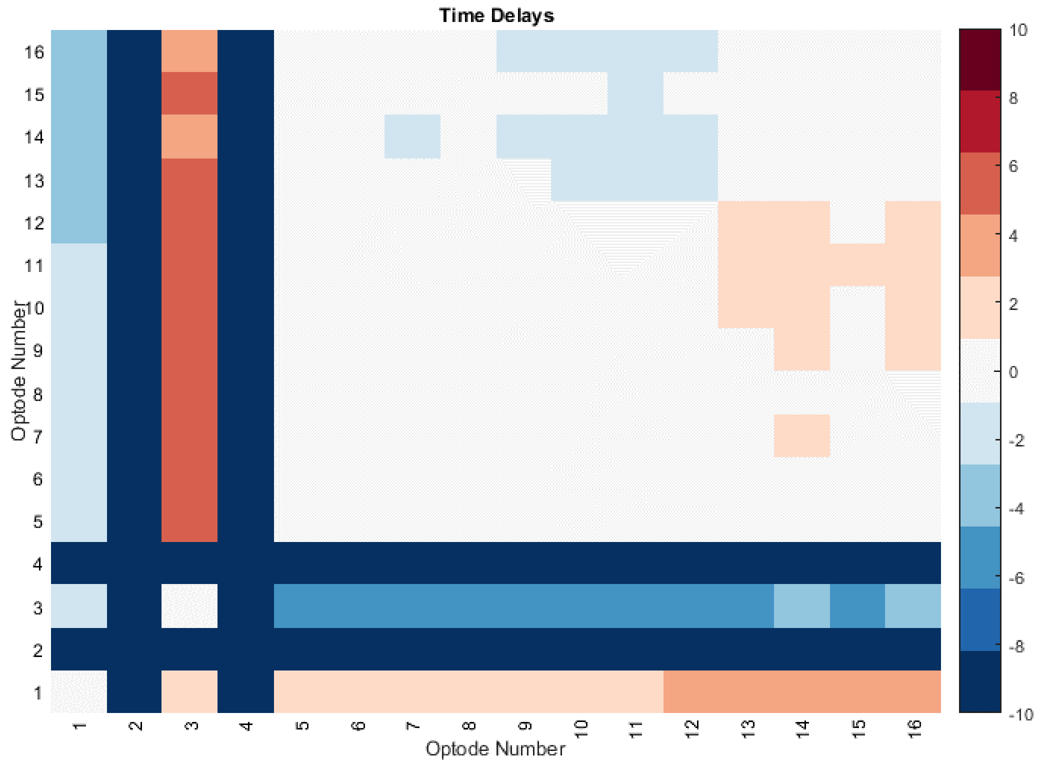
We apply the Fourier series inner product technique and find two sixteen by sixteen matrices which we represent in the form of heat plots. The values of the correlation values are provided below.



Note that the scale of the correlation heat map has a range of r = 0.55 to r = 1. Therefore, even the lowest correlation values are still moderately correlated. One quick general feature of these correlation plots is that they must be symmetric and have 1 as every diagonal entry. The correlation plot must be symmetric because the correlation between signals f and g is equivalent to the correlation between signals g and f. The diagonal starting at the bottom left and moving to top right must be identically 1 because this measures the correlation between a signal and itself which was normalized to 1.

In the figure above, there are two vertical and two horizontal grey bands at optodes two and four. These represents signals that were not recorded correctly by the detectors. Signals that report NAN can easily be spotted on these heat maps because the correlation between the signals with itself will not be 1. Qualitatively, there is strong correlation between signals 5 through 16, while there was most likely some error in taking the oxygenation levels for signals 1 through 4. This supposition is supported by the comparatively lower correlation between these signals and the other signals as well as the lack of information from optodes 2 and 4. For signals 5 through 16, the strongest correlations come when we compare the signals that are located nearest each other which is evident by the highly correlated parts of the heat map along the diagonal; the off diagonal terms are comparatively less correlated. Intuitively, this is a reasonable result as when the optode numbers are closer together, then the physical detectors are closer together and they should thus register similar signals.

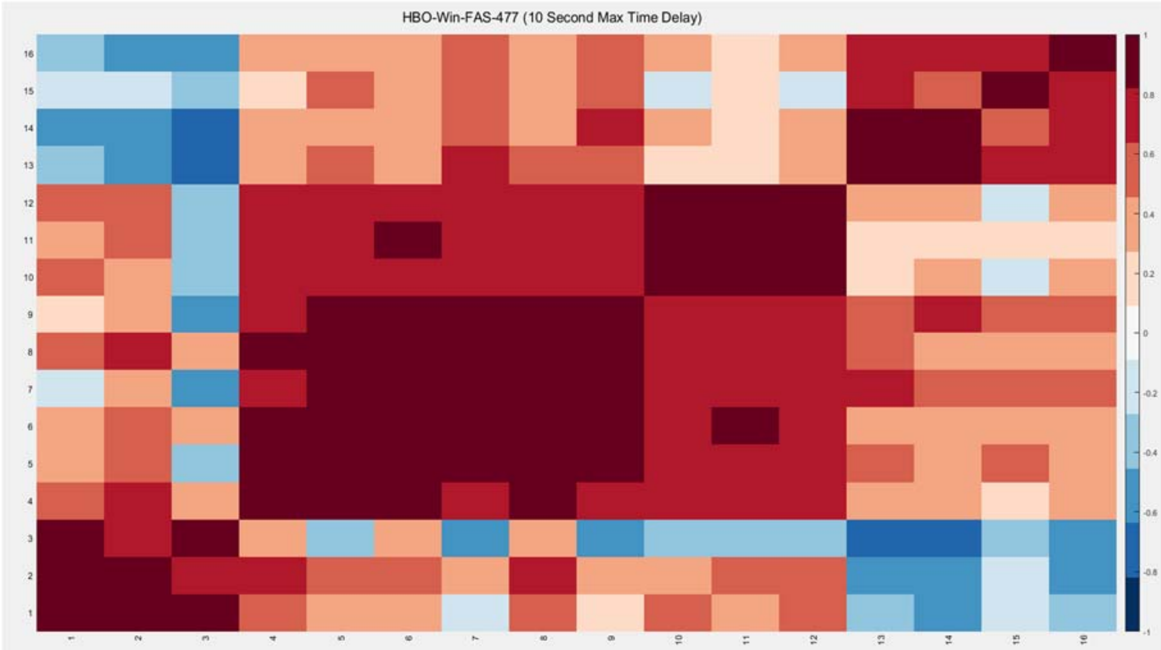
Additionally we can make a heat plot of the time delays for this subject which is reproduced below.

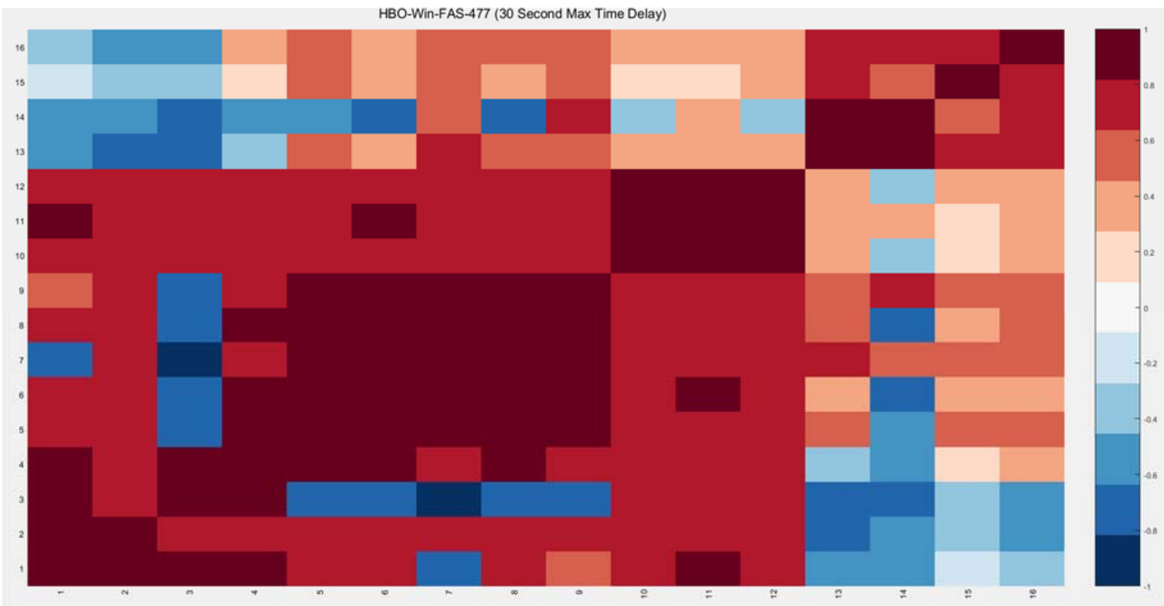


A general feature of these time delay heat maps is that they are antisymmetric and are zero along the diagonal. The former is because if the time delay between signal f and g is two seconds, then the time delay between signal g and f must be negative two seconds. The diagonal from bottom left to top right must be zero because there is no time delay between a signal and itself.

In the figure above, note that there is significantly more time delay for signals 1 through 4 than for 5 through 16. This furthers the supposition that the device did not accurately measure these signals. Additionally note that the two signals which reported NAN, signals 2 and 4, say that the time delay is negative 10 seconds. Since these detectors reported no signals, any value for the time delay is meaningless.

A second interesting feature that was found in the time delay plot is that increasing the range in which we search for time delays can significantly alter the correlation value. For instance, below are two correlation heat maps from a subject with FAS during a win trial. For the first plot, we only allowed time delays of plus or minus 10 seconds. In the second, we broadened our search and permitted time delays of plus or minus 30 seconds. The code calculates the correlation coefficient at 0.1 second intervals over the respective ranges and reports the maximum correlation value magnitude in this interval. By using the magnitude, we can account for the possibility of negative correlation between the signals. Since the first range is a subset of the second range, it follows that the magnitude of the correlation value can only increase from the first plot to the second plot, so the second heat map will appear as dark as or darker than the first.





In fact, we can clearly see a noticeable difference in the correlation values in several of the areas. From the 10 s plot, we see that signals 4 through 12 are strongly correlated, and that there is a block diagonal pattern of strong correlation. This can be seen by the rectangles of strong correlation along the diagonal. This is a reasonable result for brain behavior in that it implies that signals close to each other behave in similar fashion. For the 30 s plot, we still see the block diagonal pattern somewhat; however, signals 1 through 12 are now, including the anti-correlated signal three, all strongly correlated.

Because we found larger correlation values, it may seem logical to always expand the time delay range to the maximum possible value, which would be the length of the signal. If the difference in correlation values were always small, then we could use only the smaller range time delays.

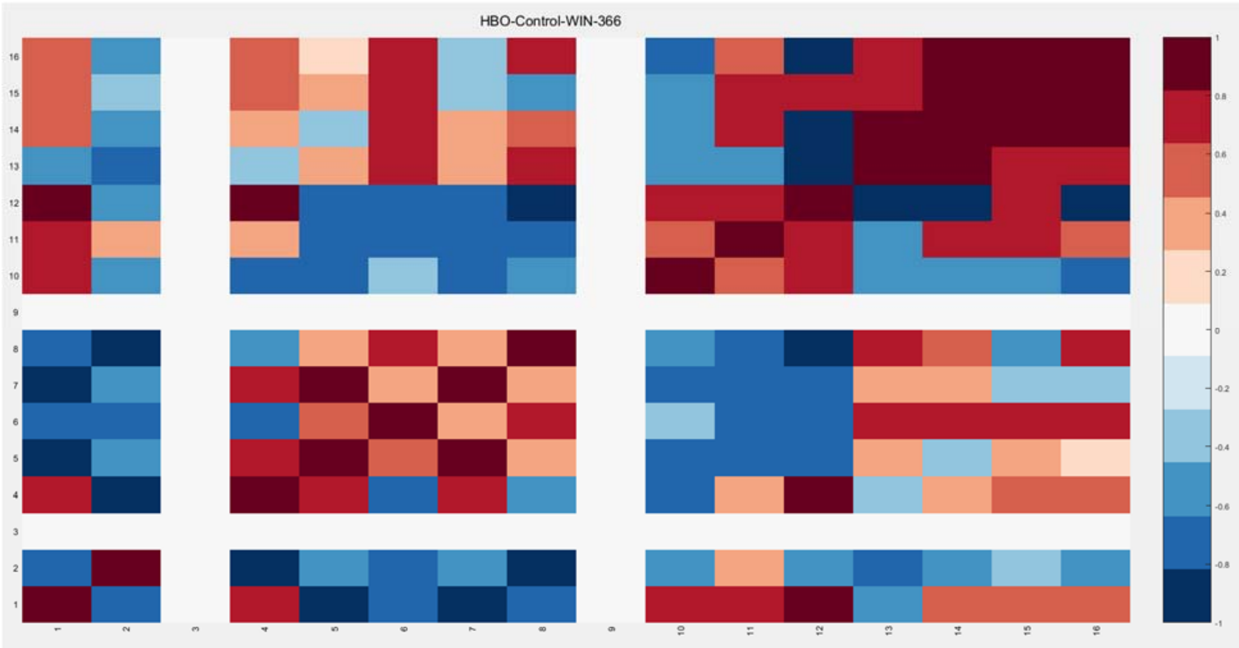
However, while we find larger correlations for the larger time delay search, these larger correlations are somewhat meaningless. To begin with, we can impose a physical constraint on how much two signals could possibly lag behind each other. 10 seconds is a reasonable estimate for the period over which a signal in one area of the brain could affect a signal in another area. Further research on oxygenation patterns in the brain may improve the accuracy of this estimate.

Moreover, recall that the Fourier series method is equivalent to the cross-correlation method, and that the cross-correlation method acts as peak finder in the sense that it matches the largest maximum of each signal with one another. This, however, does not always accurately represent how a signal changes across measurements. For instance, if one optode measures a signal with 2 peaks, the left taller than the right, and another optode measures a nearly identical signal, but with the right peak having the higher magnitude, a cross-correlation would match the largest peaks rather than matching each peak to its analogue. If the time delay is of the same order as the width of one of those peaks, however, this cannot occur. Therefore, we must impose a limit to the time delays so that the cross-correlation matches the signals in a more intuitive manner.

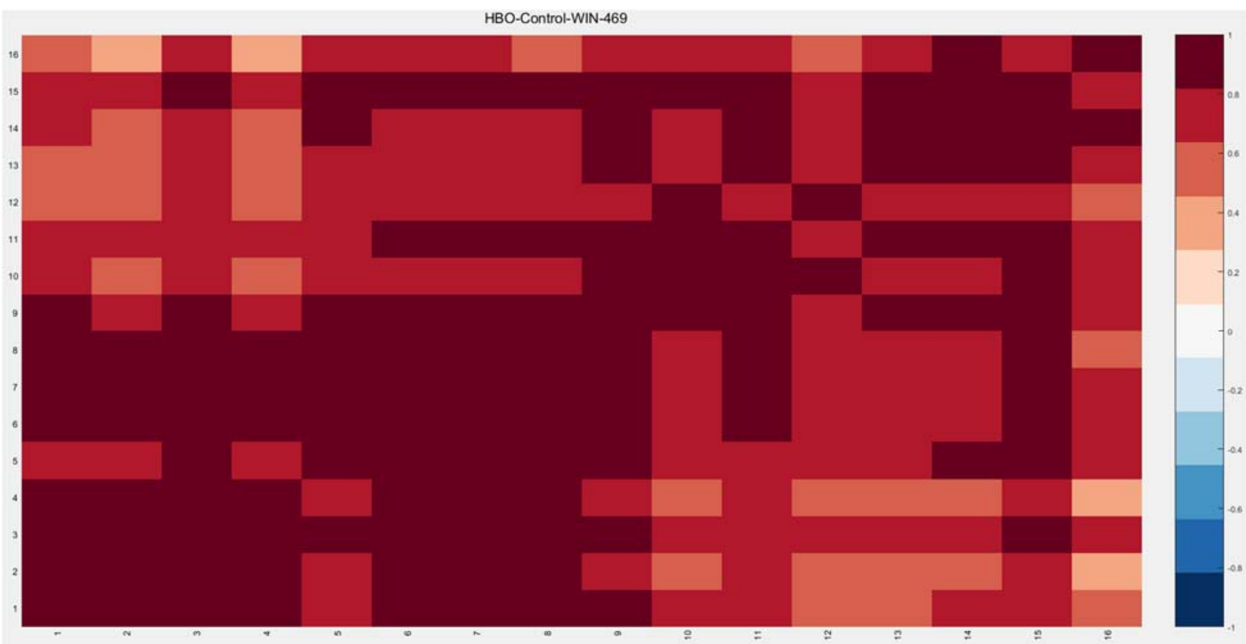
Comparing FAS and Control Correlations

The goal for this brain signals analysis is to determine if on average there is a difference between the correlations of the signals from the control group with the correlation of the signals from the FAS group. The simplest test of the strength of the correlation in the general case would be to take the magnitude of the correlation values and average them component by component. In other words, take the average of all of the 16 by 16 correlation matrices as well as the standard deviations for the control and FAS groups. In this way, we can construct a confidence interval for each component. In addition, we need to consider the error in calculating the Fourier series. For our purposes, we use a 40-term Fourier series. Since the brain signals are dominated by low frequency terms, this is sufficient to produce a standard deviation between the original signal and our approximation on the order of 10^{-3}.

While we have not been given the data to perform a full analysis, we were given some signals from subjects in the control group to test the viability of the code. The correlation of the signals from one of these subjects is reproduced below.

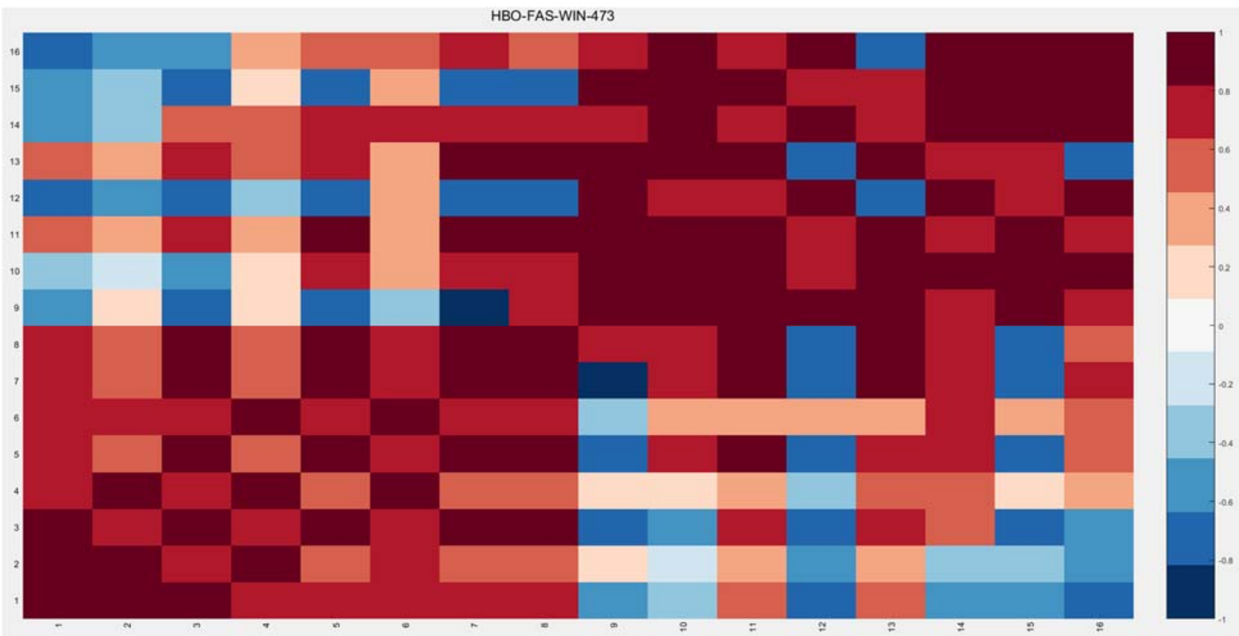


One interesting feature of this heat plot is that there is very little strong correlation which is a direct contrast to both of the previous plots shown. In fact, most of the samples that we were given had weaker correlations. We discovered that these subjects in the control group had a lower than average IQ. In fact, the plot below shows the correlations for a subject with an average IQ. In this plot, it is clear that there is strong correlation along the diagonal and weak to moderate correlation off the diagonal which represents the correlation between regions of the brain that are farther apart.



When a control group is set up, it is important for the sample to be representative of the population. It turns out that the small set of sample data that we were given to test had a disproportionate amount of subjects with below average IQ's. Interestingly, we were able to detect the lower IQ subjects using the Fourier series method which indicates that the correlation of the brain signals may be dependent upon the intellectual capability of the subject, and the difference was apparent from the Fourier series method. Of course, this dependence should be verified by a larger sample size before any conclusions can be reached.

While with our current small set of data no effect of FAS on brain signal correlation can be confirmed, it is evident that our cross-correlation plots can be successful at identifying broad neurological trends. For instance, in the plot below, which represents an FAS win trial, the signals from optodes 1 through 8 are clearly well-correlated with each other, as are signals 9 through 16, but the correlation between these two blocks is much lower. As can be seen by referencing the way the optodes were arranged relative to the brain, this shows that the two halves of the brain are communicating poorly. Whether this is an effect that is more common in subjects with FAS remains to be seen in future trials.



As a side note, this plot clearly displays a checkerboard pattern where correlation is high within the set of even or odd optode numbers but slightly lower between sets. This is an expected consequence of the setup of the detector since, as can be seen by referring again to the picture of the array, one row contains even numbers and the other row contains odd numbers. Since optodes are closer to other optodes in their row than to those in the other row, we would expect the signals to be more correlated. That this is easily observable from the cross-correlation heatmap is further confirmation that the method accurately assesses the similarity of the brain signals.

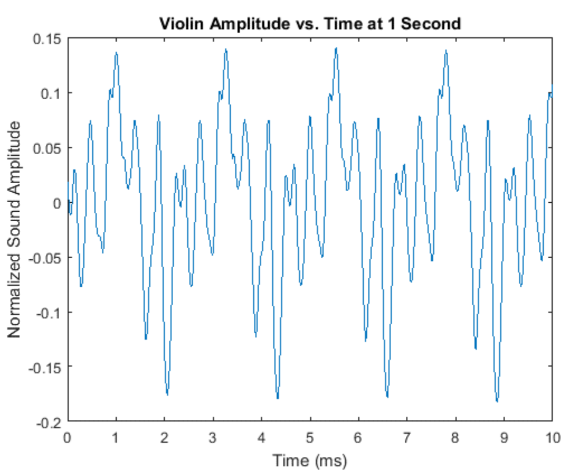
Applications to Sound Waves

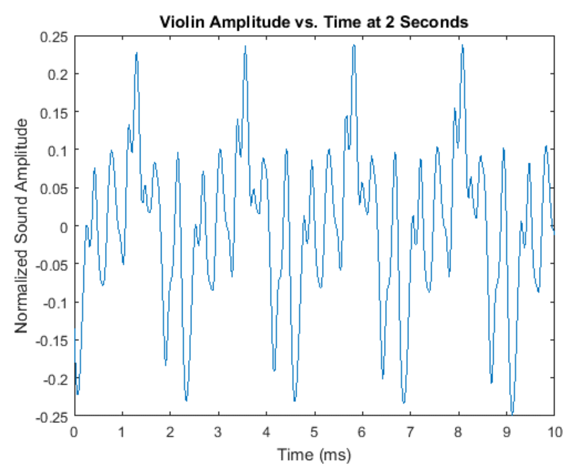
The same programs and methods that allowed us to correlate brain activity can shed light on the properties of sound waves. A natural application is analyzing the sound waves produced by musical instruments; unlike more synthetic sounds, they have nontrivial harmonics and fluctuations in amplitude and frequency that are often too subtle to hear. Even a series of 3-to-5-second audio samples of various instruments playing a sustained note contains a wealth of information that can be interpreted by cross-correlation.

Time-Dependent Correlation Loss

The unique sound produced by a musical instrument is typically described as being dependent on a similarly unique combination of harmonic amplitudes. This combination is in fact highly-time dependent, shifting constantly from second to second – while this effect is well known, our cross-correlation method provides an intuitive way to quantify the extent with which a sustained note changes its frequency distribution in time.

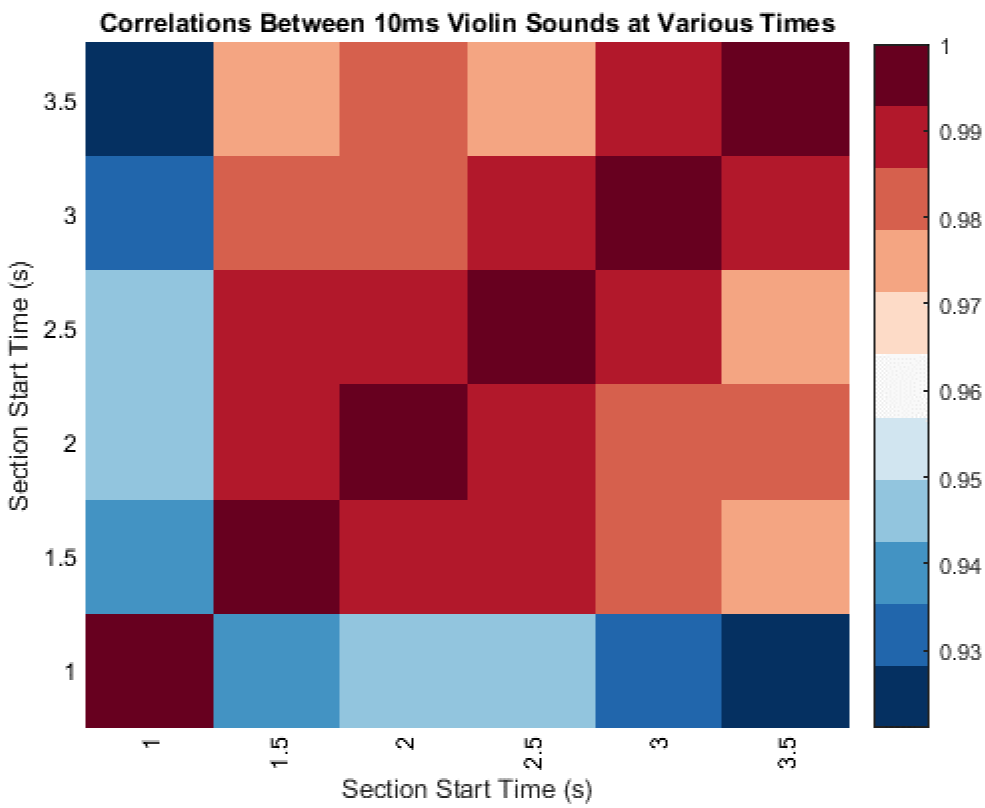
For this study, all of the instrument sound samples were playing the standard note A4, which has a frequency of 440 Hz. Thus, the dominant features of the sound wave were expected to appear in the $\frac{1\,s}{440} \sim 2.72\ ms$ range. In order to include both these features and those at lower octaves, the sections of the samples we analyzed were 10 ms long. The cross-correlation program automatically matches up waves with disparate phases, so it was not necessary to have an integer number of periods in each section or to start each section at the same point in the wave's periodic structure. To show how the time-dependent wave changes with the duration of the sustained note, we extracted 10 ms data at 0.5 s intervals for various instruments. For instance, below are plots of the amplitude (normalized with respect to the largest amplitude in the entire 5-second sample) vs. time from a violin starting at 1 second and again starting at 2 seconds:





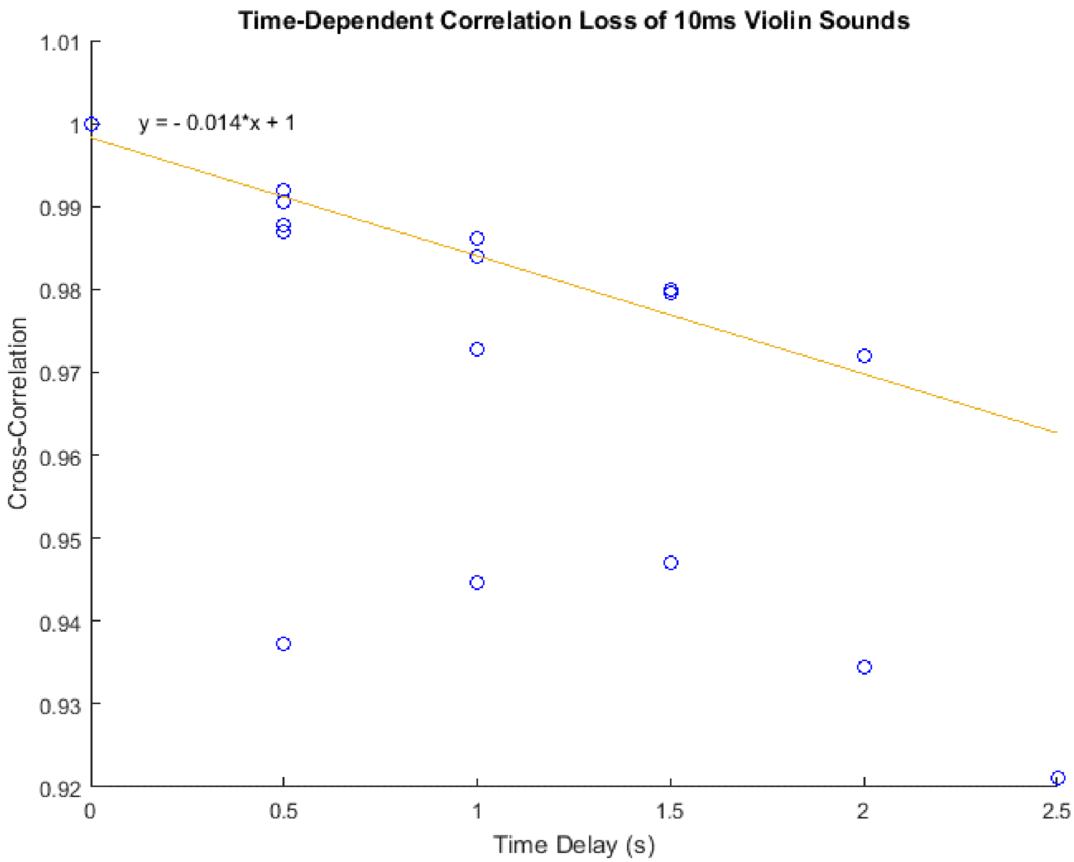
Note that the note has increased in volume from 1 second to 2 seconds, with the amplitude of the 2s wave being greater by a factor of almost 2. However, these changes are not taken into account by the correlation program since the waves are normalized such that the integral of the square of the function is 1. This is fortunate because it allows the program to discern much more subtle changes to the wave structure, such as the emergence of a small peak immediately after the largest peak of each period between 1 and 2 seconds into the audio sample.

Using the same program that correlated the 16 channels of each trial in the neuroscience study, we compared the 6 sections of the violin sample taken at each 0.5-second interval between 1s and 3.5 s. Sections at earlier and later times were omitted since the samples contained the start and end of each note, which were not the focus of this study as the changes that occur during the sustained note are much more difficult to observe by other means. Just as with the brain pattern data, the correlations between any 2 signals can be visualized as a heatmap:



Note again the plot scale: these signals are much more correlated than those from the neuroscience study, which is unsurprising because all of the signals have the same dominant frequency and, on average, the relative prevalence of harmonics characteristic of a violin. Even the first row and column, which are blue to signify that the signal changes a comparatively large amount in the third half-second of the sample, are at above 90% correlation with the rest of the sample. This sudden change is a transient effect from the start of the note at ~0.5 seconds which becomes negligible by 1.5 seconds into the sample but nevertheless differentiates the 1-second signal from the rest.

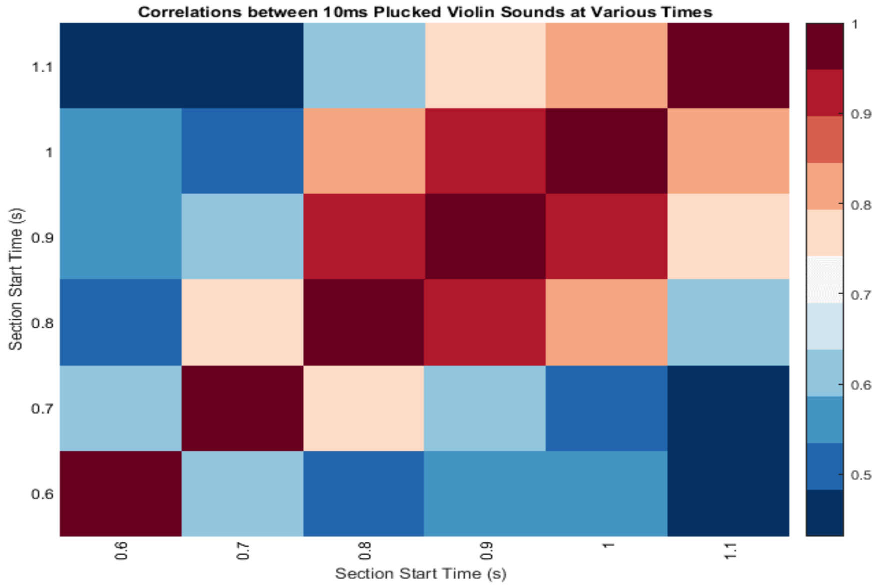
Since the frequency spectrum of the note changes with time, we would expect sections of the sample taken at longer time intervals apart to be less correlated with one another, and indeed, that is exactly what we observe, since the correlation decreases as distance from the main diagonal of the plot increases. If we graph the cross-correlation of each of this heatmap's 21 unique data points (accounting for the matrix symmetry) as a function of the time difference between them, we can clearly see a linear decrease in correlation with time once the transient effects from the 1s wave are discounted.



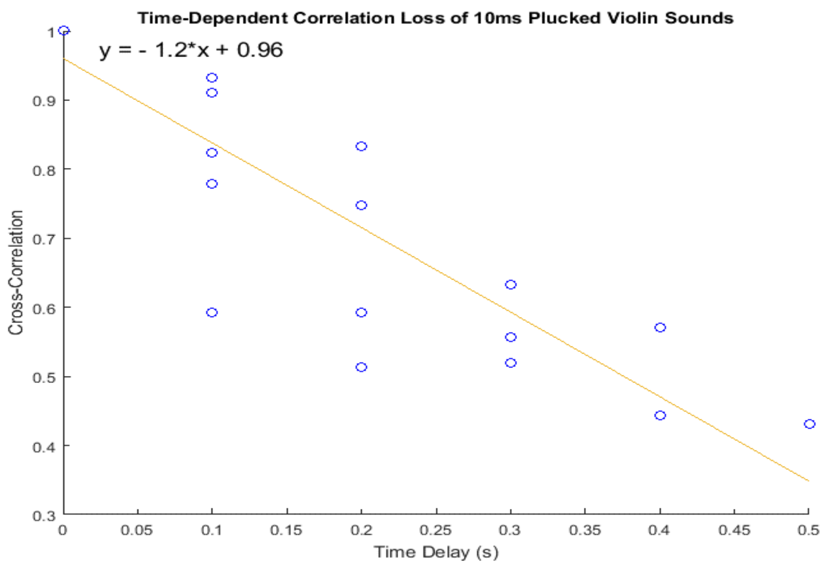
The negative slope of the line represents how fast correlation in an instrument changes, and so provides a basis for comparison between various instrument sounds. For the purposes of this paper, we can say that an A4 played on a violin has a correlation loss coefficient of 0.014 s^{-1}. However, there are too few data points to report this value with very much confidence; collecting more sound samples of instruments playing sustained notes and more data points per sample would improve the reliability of this coefficient and is therefore a good candidate for future study. Indeed, this linear change can't continue forever, so we can take it to be the first-order approximation to a more complicated relation between time delay and a loss of correlation. Further work could more precisely identify this relation.

Correlation Loss in Other Instruments

Now that these methods have been established for the violin, we can compare the rate at which the violin's sound waves become uncorrelated with the corresponding rates of other instrument sounds. One sound of particular interest is a plucked violin string. Since the sound rapidly diminishes in amplitude, it is worth asking whether it becomes uncorrelated at a similar rate. Producing a heatmap of 10 ms sounds from a plucked violin sound sample, taking data at intervals of only 0.1 seconds, shows immediately that this is the case:



When the plucked violin string's correlation is plotted as a function of time delay, it can be shown to have a correlation loss coefficient of 1.2, 2 orders of magnitude above that of the violin played with a bow.



We then analyzed an electric guitar, a clarinet, and a saxophone, which all displayed the typical time-dependent correlation loss that we observed previously in the violin and had correlation loss coefficients of 0.15, 0.027, and 0.033, respectively. These values indicate that the clarinet, saxophone, and violin produce less variable sound waves over an amount of time on the order of 1 second than an electric guitar.

While this information does not imply that frequency spectrum fluctuations of an electric guitar are substantial enough to be discernable by the human ear, further research could determine the level of correlation loss that is noticeable. For instance, a synthetic sound could be constructed to accumulate random noise over a time interval on the order of 1 second, and a participant could indicate when they notice that the sound has deviated from a constant pitch. With enough data, such a study could allow analyses like the one described in this paper to provide practical data on the qualities of a musical instrument's sound.

Direct Instrument Comparison

In addition to measuring the correlations between different sections of a single sound sample, our method can directly compare the wave structures of any two musical instruments' sustained sounds. This effectively allows us to obtain a simple value from 0 to 1 representing how much two instruments sound alike.

Due to the general rule that correlation is lost as time delay increases, a 10ms section from the middle of a sample, having the least average absolute time delay with the other signals in that sample, thus has the highest average correlation with the rest of the sample. For this reason, we selected the sound wave from 2 seconds into each sample (or 0.8 seconds for the case of the plucked violin string) to represent the sound of that instrument. We then correlated each sound wave with those of the other instruments to produce the correlation values in the table below:

	Clarinet	Saxophone	E. Guitar	P. Violin	B. Violin
Clarinet	1.0000	0.5481	0.6160	0.6527	0.5477
Saxophone	0.5481	1.0000	0.3475	0.5713	0.4132
E. Guitar	0.6160	0.3475	1.0000	0.3907	0.3991
P. Violin	0.6527	0.5713	0.3907	1.0000	0.5331
B. Violin	0.5477	0.4132	0.3991	0.5331	1.0000

This yields some surprising results. For instance, since the clarinet is in its construction more similar to the saxophone than to any of the other instruments, one would expect its sound to have a similar quality. However, the correlation values would suggest that it the time-dependent function is actually more similar to that of the electric guitar. Similarly, the plucked violin string and the bowed string do not produce very correlated sounds despite being produced by the same instrument.

This last effect may be explained by the fact that, just as we ignored the transient effect from the beginning of the bowed violin's sound sample due to the cross-correlation loss not being representative of the sustained sound, so too is the plucked violin string data unreliable. Because the correlation is lost so rapidly in the plucked string case, transient effects from the initial impact of the string dominate the sample. In this sense it may not be meaningful to compare the plucked string to the other instrument sounds at all.

Possible Future Work

This work is primarily a proof-of-concept for using normalized cross-correlations to compare signals to each other and to themselves at different times. As such, there are a number of possible extensions and applications of the ideas of this paper.

For instance, an obvious continuation of this work is to collect and analyze more instrument samples and compare them to each other using an array of cross-correlations. All of the instrument amplitude signals analyzed in this study had cross-correlations of between 0.39 and 0.66, which is a range that one would expect to see in sounds that have the same dominant frequency but are otherwise entirely unrelated (in contrast to the correlations on the order of 0.9 that were obtained when the samples were compared with themselves at later times). It would be interesting to see examples of instruments that produce much more similar signals, and to see whether this similarity in sound waves corresponds to any other connections between the instruments, such as the material from which they were made or the type of vibrations they produce (i.e. are they woodwinds, strings, percussion, etc.).

The time-dependent correlation loss study could benefit from the same type of expansion, with more instruments being analyzed to see if any connections can be drawn between the extent of the correlation loss and any attributes of the instrument. The electric guitar had significantly more correlation loss than the other instruments tested (again ignoring correlation loss due to transient effects from the start of the note). It is still unclear whether this is a trait unique to the electric guitar, whether it is a trait all electric instruments share, or whether, with more data points taken and more electric guitar sound samples analyzed, this effect will turn out to have been statistical error.

Finally, due to the sensitivity of cross-correlations to small changes in signals, this method can be used to evaluate the effects that making small changes to a musical instruments has on their sound quality. For instance, it can compare the sounds that a clarinet produces with one type of reed or mouthpiece to a competitor's, or how the position at which a string is plucked changes the sound that is produced. The number of parameters that can be manipulated and compared using cross-correlation is innumerable.

Conclusion

Cross-correlations, and, by extension, the Fourier series approximations to which they are related, can be effective methods for comparing signals, whether they are produced by musical instruments or by the brain. Their ability to produce quantities that condense the complexity of analyzing both signals to a number from 0 to 1 is both useful for interpreting the data and for understanding it in an intuitive way. Thus, the applications described in this paper represent a small fraction of the possible uses for this method.